

ОСОБЛИВОСТІ АНАЛІЗУ ТОНАЛЬНОСТІ ТЕКСТУ НА ОСНОВІ АСПЕКТІВ ПРИ ВДОСКОНАЛЕННІ КОМП'ЮТЕРНИХ СИСТЕМ

UDC 004.89

R. Havan, V. Yatsyshyn

(Ternopil Ivan Puluj National Technical University, Ukraine)

FEATURES OF SENTIMENT ASPECT-BASED ANALYSIS IN THE PROCESS OF COMPUTER SYSTEMS MODERNIZATION

Задача аналізу текстової інформації може розв'язуватись багатьма способами із застосуванням різних методів і підходів. Для оцінювання тональності текстових повідомлень у процесі вдосконалення комп'ютерних систем найбільш дієвим, з точки зору точності класифікації щодо приналежності відгуків до позитивного, негативного чи нейтрального класів, є аспектний підхід. Постановка задачі визначення тональності тексту за аспектами передбачає, що текст містить у собі оцінку деякого об'єкту. При цьому об'єкт розглядається як композиція його складових чи атрибутів, які можуть бути оцінені автором.

Зазвичай, враження про використання комп'ютерних систем, представляються у тематичних блогах або соціальних мережах, де основною одиницею тексту є відгуки користувачів, тобто текст є невеликим за об'ємом. У цьому випадку ефективним є використання підходу аспектного оцінювання думок. Аспекти можуть групуватися у категорії, для прикладу, інтерфейс користувача комп'ютерної системи, функціональність системи, поведінка системи тощо. Окрім цього, можлива наявність відгуків, де комп'ютерна система описується як єдиний об'єкт. Важливим при визначенні тональності текстової інформації є аналіз слів та виразів, які описують аспект сутності, так звані аспектні терміни.

Отже, задача аналізу текстової інформації у процесі вдосконалення комп'ютерних систем може бути поділена на ряд підзадач:

- визначення аспектних термінів;
- класифікація визначених термінів за аспектними категоріями;
- автоматичне визначення приналежності аспектів за категоріями.

Для виявлення аспектних термінів для конкретної предметної області, де використовується комп'ютерна система, можливе на основі декількох підходів.

Найбільш часто використовуваним типом аспектних термінів є явно задані терміни, які описують конкретний об'єкт, його частини або властивості, наприклад, візуальний дизайн, конкретну функцію системи, зручність використання системи та ін. Такі терміни найбільш часто виражаються іменниками або групами іменників, однак у деяких випадках, можливий опис явних аспектів за допомогою дієслів.

Іншим типом аспектних термінів є так звані неявні терміни, які представляються словами з явно вираженими оцінками компонента, які вказують на об'єкт відгуку. Зазвичай вони представляються прикметниками, наприклад, зручний (позитивний+ інтерфейс), швидкий (позитивний+відгук системи). Неявні аспектні терміни можуть комбінуватись з так званими операторами порівняння, які або змінюють або підсилюють значення оцінки щодо об'єкта або його властивості. Зокрема, це стосується застосування частки «не» (не дуже зручний, не надто швидкий). Важливість таких аспектних термінів для словників автоматичних систем аналізу тональності текстової інформації полягає в тому, що у ситуаціях неможливості розпізнавання термінів, що можуть бути пов'язані з помилками при наборі тексту, вживанні специфічних переключених термінів та ін. факторів неявні аспектні терміни дають змогу визначити ставлення автора по відношенню до деякої аспектної категорії.